

Quinary lattice model of secondary structures of proteins

S.V.Kozyrev*, I.V.Volovich†

March 8, 2013

Abstract

We construct a model of a lattice polymer which describes secondary structures of proteins. In this model the energy of a conformation of a polymer is equal to a sum of energies of conformations of segments of the polymer chain of the length five.

We show that for this model with cooperative interaction all conformations with minimal energy are combinations of lattice models of alpha-helix and beta-strand. We show that for lattice polymers of the length not longer than 38 monomers we can describe all conformations with minimal energy.

1 Introduction

In the present paper we construct the model of lattice polymer (the quinary lattice model) where the minima of energy will have the form of lattice models of combinations of alpha-helices and beta-strands. The energy of a protein in this model will be equal to the sum over conformations of segments of a polymer of the length five. Therefore instead of consideration of the explicit interaction between amino acids in a protein we consider a cooperative interaction in a polymer chain.

Lattice models of polymers were extensively discussed in the literature, cf. [1, 2, 3, 4]. In [5, 6, 7] design of sequences of polymers which fold to

*Steklov Mathematical Institute, Russian Academy of Sciences

†Steklov Mathematical Institute, Russian Academy of Sciences

given conformations was discussed. For the review of physics of proteins see [8, 9]. A protein globule was considered as a crumpled globule [10], in [11] the relation of this kind of globules and random hierarchical networks was discussed.

The problem of protein folding, i.e. folding to the native of a protein instead of performing search over a vast space of conformations is well known. The approach of folding funnels based on the idea of cooperative (collective) interactions in proteins was discussed, see for example [12].

In [13] the review of application of combinatorial algorithms to protein folding was given and investigation of complexity of these algorithms was discussed.

In [14] the GOR method of analysis of conformations of proteins was developed. In this approach the probability of an amino acid to belong to some secondary structure was considered as dependent on a short segment of the protein containing this amino acid.

In [15] proteins were considered as combinations of short sequences (in particular 5-tuples) of amino acids, in [16] the statistics of conformations of short segments of proteins was discussed.

In [17] native states of proteins were described with the help of combinations of solitons.

The structure of the present paper is as follows.

In section 2 we introduce the quinary lattice model of a polymer and show that the lattice models of alpha-helix and beta-strand are the energy minima for the introduced model.

In section 3 we describe minima of energy of the proposed model and show that these conformations can be considered as combinations of lattice alpha-helices and beta-strands.

In section 4 we consider the heteropolymer version of the introduced model and show that in this case we can describe polymers which possess a native tertiary structure — a conformation which minimizes the energy and is uniquely defined by the sequence of the polymer. The obtained tertiary structure is a combination of secondary structures described in sections 2 and 3.

In section 5 we discuss the model of energy of a lattice polymer which is a combination of the quinary lattice model and the standard model of the nearest (in the lattice) neighbors interaction of amino acids.

2 The quinary lattice model

The standard model of energy of lattice polymers has the following form, see [1], [2]. One considers a linear lattice polymer (a finite sequence of monomers connected by edges of the length one), the monomers are situated at vertices of the cubic lattice \mathbb{Z}^3 . A conformation of a polymer of the length N is a sequence of neighbor vertices without self-intersections in the cubic lattice \mathbb{Z}^3 , i.e. the injective map

$$\Gamma : \{1, \dots, N\} \rightarrow \mathbb{Z}^3, \quad (1)$$

where neighbor natural numbers map to neighbor (i.e. distance one) vertices of the lattice \mathbb{Z}^3 . In the following we will denote Γ also the image of this map.

Energy of the conformation Γ in the standard model is proportional to the following sum

$$E_2(\Gamma) = - \sum_{1 \leq i < j \leq N} \delta(d(\Gamma(i), \Gamma(j))), \quad (2)$$

where $\Gamma(i)$ is the i -th monomer in the polymer, $d(\cdot, \cdot)$ is the distance in \mathbb{Z}^3 , $\delta(1) = 1$, $\delta(i) = 0$, $i > 1$. Non zero contributions to this sum come from the contacts of pairs of monomers in the conformation Γ .

The above model takes into account only the number of contacts in the conformation but does not distinguish between the conformations with the different geometry. Therefore this model does not describe formation of secondary structures (special preferred conformations of a protein), in particular, alpha-helices and beta-strands.

In the present paper we introduce the following model of lattice polymer. The energy of a polymer in this model will be equal to the sum of contributions where each of the contributions depends on conformation of a segment of a lattice polymer of the length 5 (i.e. which contains five monomers). Therefore the energy of a polymer of the length N in conformation Γ has the form

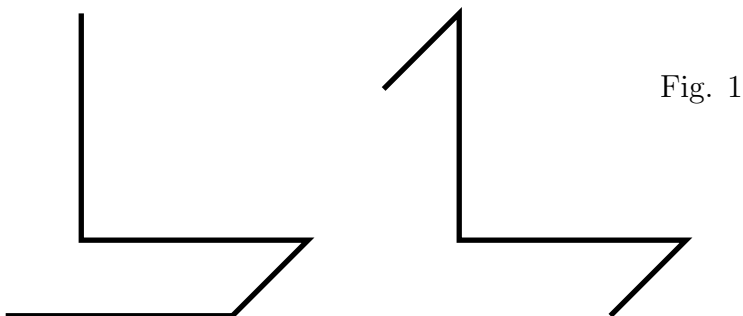
$$E_5(\Gamma) = - \sum_{i=3}^{N-2} \Phi(\Gamma_i). \quad (3)$$

Here Γ_i is a conformation of the i -th 5-tuple of monomers in a polymer, i.e.

$$\Gamma_i = (\Gamma(i-2), \Gamma(i-1), \Gamma(i), \Gamma(i+1), \Gamma(i+2)), \quad i = 3, \dots, N-2,$$

Conformation 1

Conformation 2



where $\Gamma(i)$ is the i -th monomer in a polymer of the length N , Φ is some function of conformations of 5-tuples of monomers. The function Φ is taken to be invariant with respect to lattice rotations and translations of conformations of 5-tuples.

The intersection of the two neighbor 5-tuples Γ_i and Γ_{i+1} contains the four monomers $\Gamma(i-1), \Gamma(i), \Gamma(i+1), \Gamma(i+2)$.

It is easy to see that (modulo lattice translations and rotations) there exist 30 different conformations of 5-tuples without self-intersections. We choose the function Φ as follows — $\Phi(\Gamma)$ is equal to zero for all conformations except the conformations denoted 1 and 2 for which $\Phi(1) = \Phi(2) = 1$, cf. Fig. 1 (by definition Φ will be also equal to one for lattice translations and rotations of conformations 1 and 2).

At Fig. 1 the edges connect neighbor monomers, thus a segment of a polymer containing five monomers will contain four edges. Conformations 1 and 2 are segments of a right handed helix, see. Fig. 2.

Proof of the following lemma is straightforward.

Lemma 1 *Conformation of any 5-tuple of neighbor monomers in periodic conformations of lattice polymers at Fig. 2 is either conformation 1 or conformation 2 of Fig. 1. Therefore these periodic conformations are minima of energy (3).*

The left conformation at Fig. 2 may be considered as a lattice model of α -helix, the right conformation can be considered as a model of β -strand. Therefore in the model under consideration we observe the two most important examples of secondary structures.

In real proteins β -strands are stabilized by hydrogen bonds between parallel chains, therefore to have a stable conformation one needs at least two

α -helix

β -strand

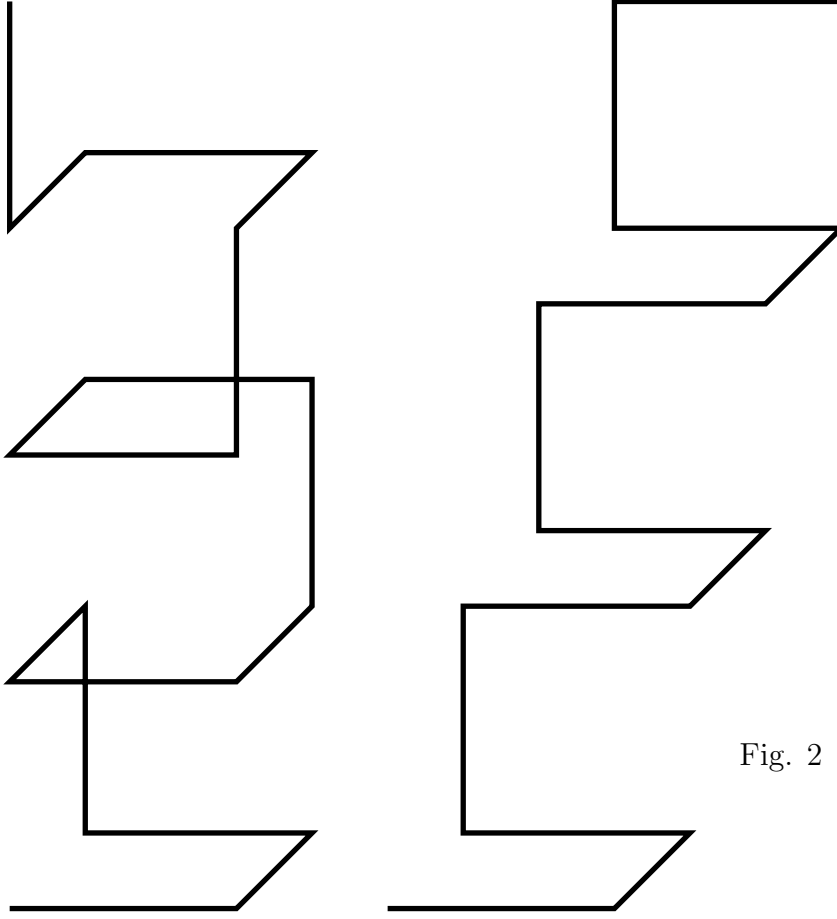


Fig. 2

parallel β -strands. Our model neglects hydrogen bonds and considers conformations which model folds of peptide chain in secondary structures as energetically profitable. Therefore this model can not pretend to give a realistic description of conformations of real proteins. Our aim is to describe qualitatively the effect of existence of secondary structures.

In the next section we will show that the described lattice α and β conformations and their combinations are all possible minima of energy for the quinary lattice polymer model (3).

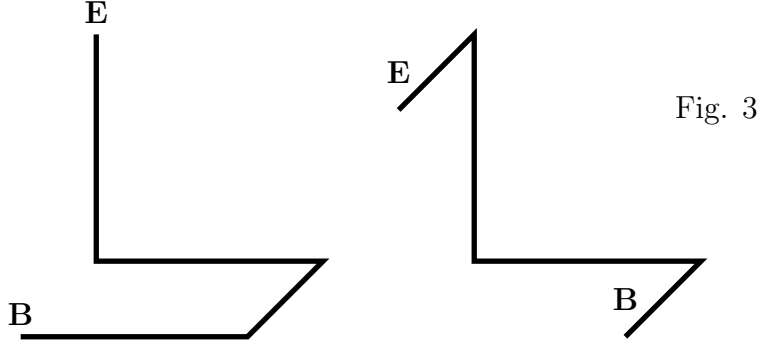


Fig. 3

3 Enumeration of minimal conformations

Let us fix the directions in the 5-tuples of monomers at Fig. 1. We choose the beginnings and the ends of 5-tuples, see Fig. 3 (where B denotes the beginning and E denotes the end). The obtained conformations of directed polymers we denote $\vec{1}$ and $\vec{2}$ (the left and the right conformations at Fig. 3 correspondingly). The same conformations but with the opposite directions (from the end to the beginning) we denote $\overleftarrow{1}$ and $\overleftarrow{2}$.

Conformations of lattice polymers which are minima of the energy (3), i.e. conformations of lattice polymers for which any 5-tuple of neighbor monomers has one of the conformations $\vec{1}$, $\vec{2}$, $\overleftarrow{1}$, $\overleftarrow{2}$ are called *minimal*. Note that the neighbor 5-tuples of monomers in a polymer (or 4-tuples of edges shown at Fig. 3) have the intersection containing 4 monomers (or 3 edges). We consider these conformations as models of secondary structures in proteins.

A lattice polymer in a minimal conformation Γ generates a sequence $\Gamma_3\Gamma_4\dots\Gamma_{N-2}$ of conformations of 5-tuples of monomers (when we read the sequence of monomers in the polymer from the beginning to the end), where $\Gamma_i \in \{\vec{1}, \vec{2}, \overleftarrow{1}, \overleftarrow{2}\}$. The conformation Γ of a polymer can be restored from the sequence $\Gamma_3\Gamma_4\dots\Gamma_{N-2}$ of conformations of 5-tuples.

We say that a minimal conformation $\Gamma_3\Gamma_4\dots\Gamma_{M-2}$ is a continuation of a minimal conformation $\Gamma_3\Gamma_4\dots\Gamma_{N-2}$, $N < M$ if the second sequence of symbols is a segment of the first sequence (i.e. the first sequence is obtained from the second by adding some symbols in the beginning and the end).

Which sequences $\Gamma_3\Gamma_4\dots\Gamma_{N-2}$ of the conformations $\vec{1}$, $\vec{2}$, $\overleftarrow{1}$, $\overleftarrow{2}$ can be generated by minimal conformations of a lattice polymer? It is not always possible to combine a couple of conformations $\vec{1}$, $\vec{2}$, $\overleftarrow{1}$, $\overleftarrow{2}$ of 5-tuples into a single conformation of a 6-tuple due to geometric restrictions, see the next

lemma.

Lemma 2 1) Possible pairs of neighbor conformations of 5-tuples in the sequence $\Gamma_3\Gamma_4\ldots\Gamma_{N-2}$ related to some minimal conformation of a lattice polymer are described by the following table

	$\vec{1}$	$\overleftarrow{1}$	$\vec{2}$	$\overleftarrow{2}$
$\vec{1}$	−	+	+	−
$\overleftarrow{1}$	+	−	−	−
$\vec{2}$	−	−	−	+
$\overleftarrow{2}$	−	+	+	−

(i.e. for any pair of symbols denoted by + in the table above there exists a minimal conformation of a lattice polymer of length 6).

2) Any conformation described by a triple of symbols from $\{\vec{1}, \vec{2}, \overleftarrow{1}, \overleftarrow{2}\}$ permitted by the above table corresponds to some minimal conformation of a lattice polymer of length 7 except the triples $\vec{2}\overleftarrow{2}\vec{2}, \overleftarrow{2}\vec{2}\overleftarrow{2}$.

Theorem 4 below shows that the above lemma describes all possible geometric restrictions for lattice polymers of the length shorter or equal to 38, i.e. for any sequence of conformations of 5-tuples $\Gamma_3\Gamma_4\ldots\Gamma_{N-2}$, $N \leq 38$, which satisfies the conditions of the above lemma 2, there exists the corresponding minimal conformation (without self-intersections) of a lattice polymer. Therefore in order to construct minimal conformations of sufficiently short lattice polymers it is sufficient to take into account the geometric restrictions for neighbor couples and triples of 5-tuples Γ_i .

Let us consider the following periodic sequences of conformations of 5-tuples satisfying the conditions of lemma 2

$$(\vec{1}\vec{2}\overleftarrow{2}\overleftarrow{1})\ldots, \quad (\vec{1}\overleftarrow{1})\ldots$$

Here the periods are shown in brackets (i.e. one can iterate the sequence in brackets). The corresponding conformations of lattice polymers are shown at Fig. 4.

The next lemma describes the sequences of conformations of 5-tuples for the lattice alpha and beta structures described in lemma 1.

Conformation $\alpha = \overrightarrow{1} \overrightarrow{2} \overleftarrow{2} \overleftarrow{1}$

Conformation $\beta = \overrightarrow{1} \overleftarrow{1}$

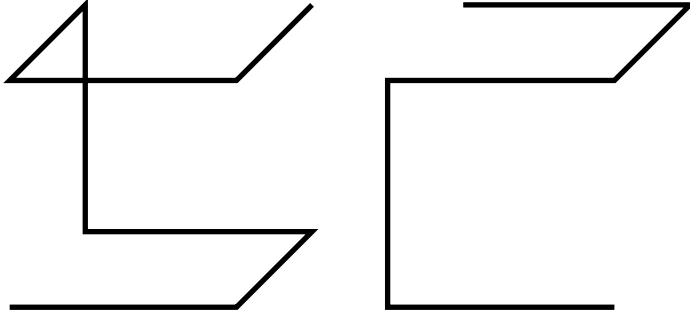


Fig. 4

Lemma 3 *Periodic sequences of conformations of 5-tuples with the periods*

$$\alpha = (\overrightarrow{1} \overrightarrow{2} \overleftarrow{2} \overleftarrow{1}), \quad \beta = (\overrightarrow{1} \overleftarrow{1}),$$

see Fig. 4, correspond to conformations of a lattice polymer without self-intersections. Moreover these conformations are lattice α -helix and β -strand (see Fig. 2) correspondingly.

The next theorem describes all minimal conformations for the model (3). We show that all these conformations correspond to combinations of α and β structures and for polymers not longer than 38 all such combinations are possible (correspond to conformations without self-intersections). For longer polymers self-intersections are possible therefore only part of combinations of α and β structures correspond to minimal conformations of lattice polymers.

Theorem 4 1) *Any minimal conformation of a lattice polymer (3) with the length $N > 6$ has the following form:*

The corresponding sequence $\Gamma_3 \Gamma_4 \dots \Gamma_{N-2}$ of conformations of 5-tuples can be obtained from some sequence of α and β structures, $\alpha = \overrightarrow{1} \overrightarrow{2} \overleftarrow{2} \overleftarrow{1}$, $\beta = \overrightarrow{1} \overleftarrow{1}$ by elimination of a finite number of symbols $\overrightarrow{1}$, $\overrightarrow{2}$, $\overleftarrow{1}$, $\overleftarrow{2}$ in the beginning and the end of the sequence.

2) *All conformations of a lattice polymer with the length $6 < N < 39$ obtained as above do not contain self intersections.*

There exists a sequence of conformations of 5-tuples corresponding to the conformation of a lattice polymer of the length 39 with self-intersections.

Proof By the definition a minimal conformation of lattice polymer of the length N generates a sequence $\Gamma_3 \Gamma_4 \dots \Gamma_{N-2}$ of conformations of 5-tuples of monomers, $\Gamma_i \in \{\overrightarrow{1}, \overleftarrow{1}, \overrightarrow{2}, \overleftarrow{2}\}$.

Let us discuss how the conformations 2 (i.e. $\overleftarrow{2}$ or $\overrightarrow{2}$) can be situated in this sequence. By lemma 2 we can not have more that two consecutive

Conformation $\alpha = \overrightarrow{1} \overrightarrow{2} \overleftarrow{2} \overleftarrow{1}$

Conformation $\beta = \overrightarrow{1} \overleftarrow{1}$

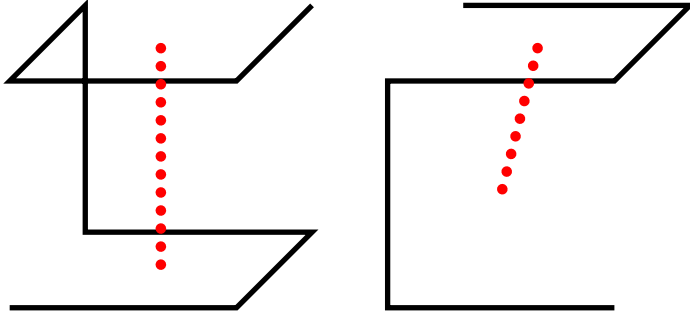


Fig. 5

symbols 2 in the sequence. A single conformation 2 can be situated either at the beginning or the end of the sequence. If the polymer is longer than six monomers the minimal conformation of this polymer can not contain a subsequence $\overleftarrow{2} \overrightarrow{2}$ (since this sequence can be continued only by 2 which is forbidden).

Therefore conformations 2 can be found inside a sequence corresponding to a minimal conformation only in pairs and for polymers longer than six monomers this pair has the form $\overrightarrow{2} \overleftarrow{2}$. Any pair of this form by lemma 2 should be augmented by conformations 1 (again modulo the boundaries of the sequence), and the corresponding conformation (which corresponds to a segment of a lattice polymer of the length eight) will have the form $\alpha = \overrightarrow{1} \overrightarrow{2} \overleftarrow{2} \overleftarrow{1}$.

The part of the sequence of conformations of 5-tuples of monomers which does not contain conformations 2 contains iterations of the conformation $\beta = \overrightarrow{1} \overleftarrow{1}$. This implies the first statement of the theorem.

In order to prove the second statement of the theorem let us consider Fig. 10 with a lattice polymer of the length 38 in the minimal conformation $\alpha\beta\alpha\alpha\beta\alpha\alpha\beta\alpha\alpha\beta$. One can see that any minimal conformation which is a continuation of this conformation will have self-intersections.

We have already checked that iterations of α or β structures are minimal conformations without self-intersections. Let us prove that minimal conformations of sufficiently short lattice polymers containing a mixture of α and β structures also do not contain self-intersections.

Let us put in correspondence to α and β structures their *central lines* as shown at Fig. 5 by dotted lines. For the α -structure the central line connects the centers of the opposite faces of the cube at Fig. 5 (edges of the α -structure will be the edges of this cube). Analogously for the β -structure the central line will connects the adjacent faces of the corresponding cube.

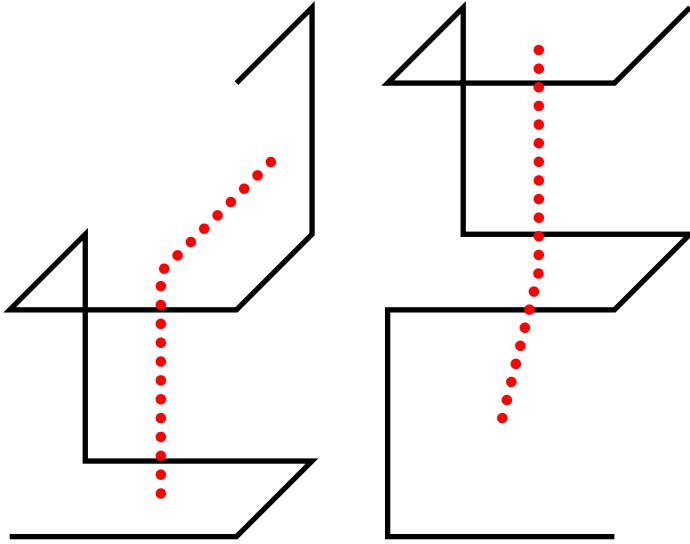
Conformation $\alpha\beta$ Conformation $\beta\alpha$ 

Fig. 6

The different combinations of α and β structures can be found at the figures 6, 7, 8, 9. One can find that:

1) Central line of a minimal conformation of a lattice polymer is a continuous broken line;

2) For the joint of two α structures or two β structures the central lines of joint structures will be parallel, therefore central lines of $\alpha\alpha\dots$ and $\beta\beta\dots$ will be straight;

3) For the joint of the different structures ($\alpha\beta$ or $\beta\alpha$) the central line breaks with the angle 135° .

For obtaining a self-intersection of a minimal conformation of a lattice polymer we need several breaks of the central line at contacts of α and β structures. The central line have to rotate for more that 180° (i.e. we need more than four breaks). One can see that the shortest lattice polymer with minimal conformation with self-intersection will correspond to some continuation of the conformation at Fig. 10, for example $\beta\beta\alpha\alpha\beta\alpha\alpha\beta\alpha\alpha\beta\vec{1}$.

Therefore any minimal conformation described at the statement of the theorem for a lattice polymer with the length not larger than 38 can be realized without self-intersections.

This finishes the proof of the theorem. \square

Remark Theorem 4 shows that for a lattice polymer with the length larger than six monomers the minima of energy (3) have the form of combinations of lattice α and β structures, and for polymers not longer than 38 all such

Conformation $\alpha\beta\alpha$

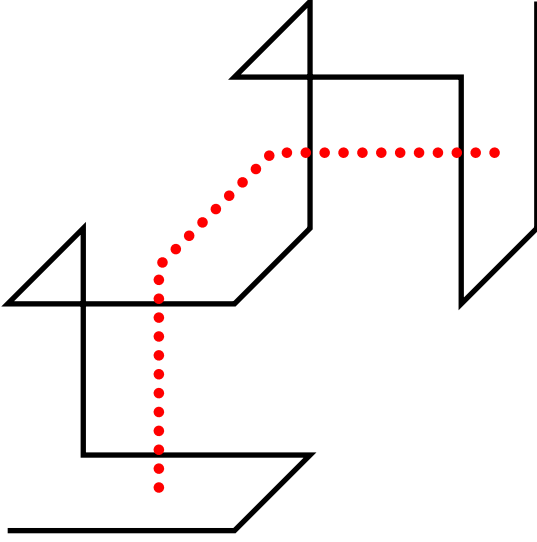


Fig. 7

combinations do not have self-intersections. Therefore collective interactions in models of lattice polymers are able to describe the formation of secondary structures in proteins. Arising of self-intersections for long lattice polymers is natural — for real proteins not all arbitrary combinations of α and β structures are possible.

4 Heteropolymers and native conformations

In the present section we consider a heteropolymer analogue of the quinary model of lattice polymer (3). We consider a lattice polymer which consists of monomers of the two kinds A and B . This polymer has the sequence of monomers S and the conformation Γ . We define the energy of the polymer by the following modification of the formula (3):

$$E_{5'}(S, \Gamma) = - \sum_{i=3}^{N-2} [\#_B(S_i) (\Phi_1(\Gamma_i) + \epsilon \Phi_2(\Gamma_i)) + \#_A(S_i) (\Phi_2(\Gamma_i) + \epsilon \Phi_1(\Gamma_i))] . \quad (4)$$

Here S_i is the sequence of monomers in the i -th 5-tuple of monomers in the polymer, Γ_i is the conformation of the i -th 5-tuple, the function Φ_1 is equal to one for the conformation 1 at Fig. 1 and to zero for all other conformations (i.e. this is a characteristic function of the conformation 1 in the space of

Conformation $\alpha\beta\beta\alpha$

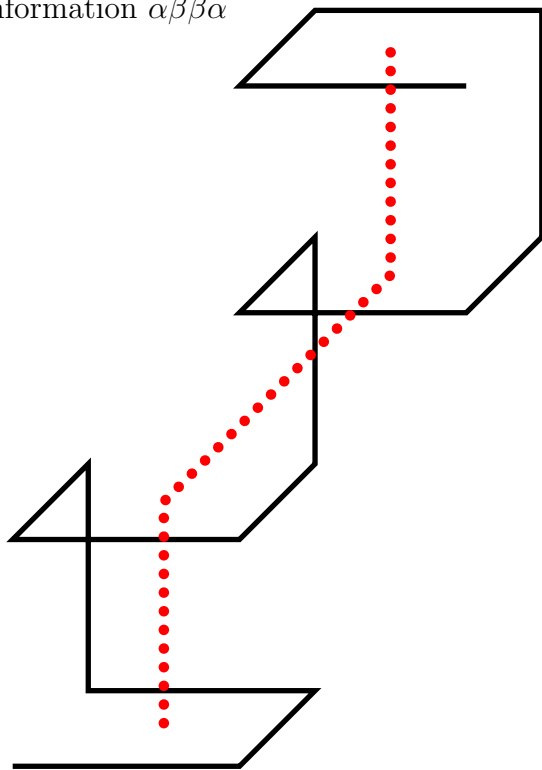


Fig. 8

Conformation $\alpha\beta\alpha\beta\alpha$

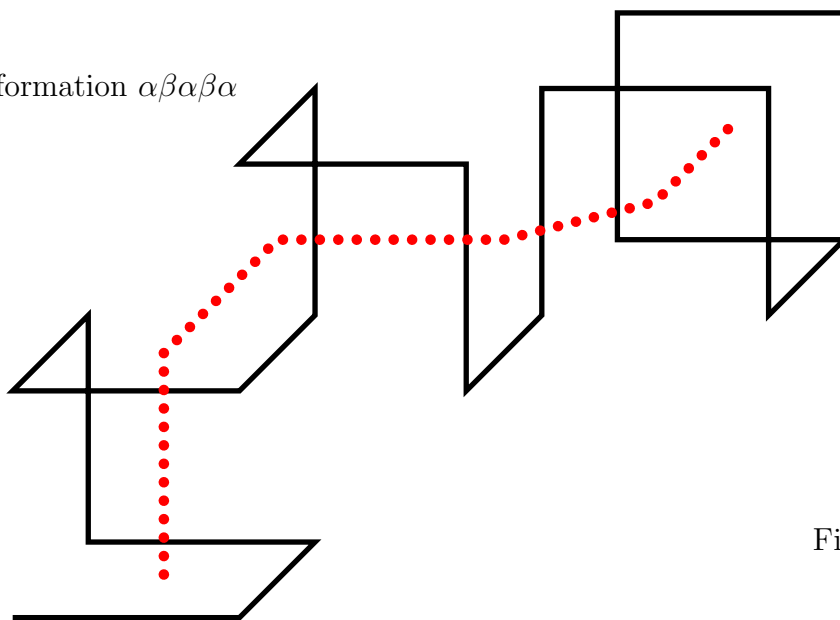


Fig. 9

Conformation $\overleftarrow{\Gamma} \beta \alpha \alpha \beta \alpha \alpha \beta \alpha \alpha \beta \overrightarrow{\Gamma}$ (38 vertices)

there exists a continuation with self-intersection

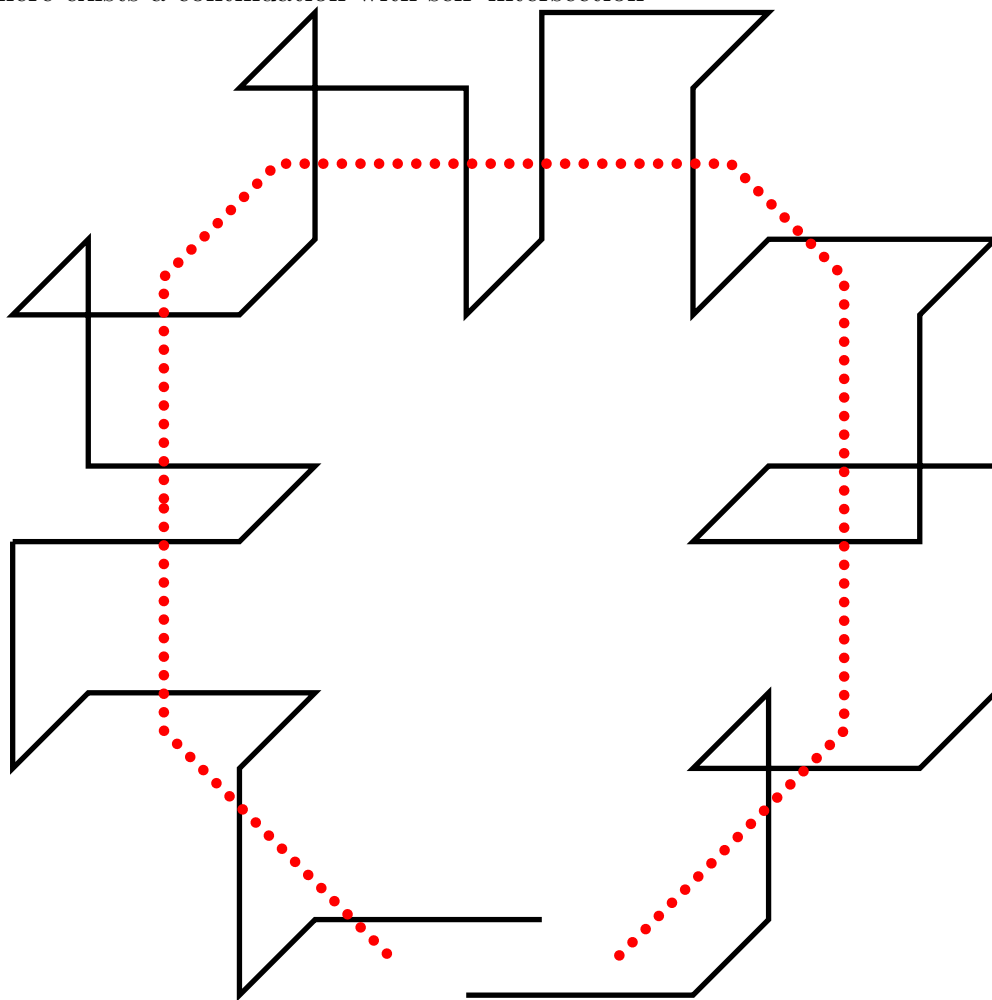


Fig. 10

conformations of 5-tuples of monomers)¹, analogously the function Φ_2 is the characteristic function of the conformation 2 at Fig. 1 (thus $\Phi = \Phi_1 + \Phi_2$ in formula (3)). The $\#_A(S_i)$ denotes the number of monomers of the kind A in the i -th 5-tuple of monomers (thus $0 \leq \#_A(S_i) \leq 5$), ϵ is a positive parameter, $0 < \epsilon < 1$. The analogous notations are used for $\#_B(S_i)$ – the number of monomers of the kind B in the i -th 5-tuple.

With the above choice of energy a lattice polymer which contains only monomers B will have the minimum of energy in the conformation equal to a single β -strand, since a β -strand contains only conformations $\overrightarrow{1}$, $\overleftarrow{1}$ of 5-tuples of monomers.

Analogously a lattice polymer which consists only of monomers A will maximize the number of conformations $\overrightarrow{2}$, $\overleftarrow{2}$ of 5-tuples in the conformation of the polymer. This can be achieved at lattice α -helix (in an α -helix a half of conformations of 5-tuples of monomers are of the type 1 and a half are of the type 2). The term $\epsilon\Phi_1(\Gamma_i)$ in expression (4) for energy makes the conformations $\overrightarrow{1}$, $\overleftarrow{1}$ more profitable energetically than a conformation which is neither of the type 1 nor of the type 2.

Therefore a lattice heteropolymer with the energy (4) for some sequences S of monomers A and B will possess a native tertiary structure — a conformation which is uniquely defined by the sequence S of monomers (the primary structure of the polymer), minimizes the energy (4) and consists of a combination of secondary structures (lattice α -helices and β -strands).

Remark We do not claim that for an arbitrary sequence S of monomers A and B for a lattice polymer with energy (4) there exists a unique native structure. Real copolymers of amino acids also possess native tertiary structures only for some sequences of amino acids.

5 Interaction of secondary structures

Let us discuss the following generalization of the introduced in the present paper model of lattice polymers. For this generalization the expression for energy of a lattice polymer is given by the following linear combination

$$E = E_5 + \lambda E_2, \quad (5)$$

¹As in section 2 we consider the conformations modulo lattice translations and rotations.

where $\lambda > 0$ is a small parameter, E_2 is given by (2) and E_5 is given by (3).

Let us consider the conformation of a lattice polymer in the form of sufficiently long lattice α -helix. If $\lambda = 0$ then this conformation will be a minimum of energy (5).

Let us consider also the conformation which has the form of two shorter parallel α -helices in contact (i.e. the distance between the helices is equal to one). This conformation may be considered as an α -helix folded in half. If λ is larger than some threshold then the energy of a folded in half α -helix will be larger than the energy of a single longer α -helix since the energetic gain will contain contributions (2) of contacts of a large number of monomers in the two α -helices, and energetic loss will come from contributions (3) of small number of 5-tuples of monomers in the area where α -helix was folded in half.

We have obtained the tertiary structure of a polymer — a compact globule built of secondary structures. Let us note that in this model it is important that the contributions to the energy of conformation from cooperative interaction (3) (which generates secondary structures) is larger than the energy of contacts of monomers in (2).

6 Conclusion

We construct a model of a lattice polymer (the quinary lattice model) which describes secondary structures of proteins. In this model the energy of a conformation of a polymer is equal to a sum of energies of conformations of segments of the polymer chain of the length five.

We show that for this model with cooperative interaction all conformations with minimal energy are combinations of lattice models of alpha-helix and beta-strand. We show that for lattice polymers of the length not longer than 38 monomers we can describe all conformations with minimal energy.

Models of energy of lattice polymers where contributions from contacts of several monomers were taken into account were discussed in the literature [2], but dependence on conformations of short segments of a polymer was not considered. The corresponding collective contributions were discussed as corrections to the contact interaction (2) of monomers. In our model (3) the cooperative interaction and conformational dependence of energy of interaction are principal.

We do not pretend that the model proposed in the present paper gives

a realistic description of conformations of proteins. Our aim was to give a qualitative demonstration of the importance of conformationally dependent cooperative interactions at short segments of a protein for creation of secondary structures.

Acknowledgments This paper was partially supported by the grants of the Russian Foundation for Basic Research RFBR 11-01-00828-a and 11-01-12114-ofi-m-2011, by the grant of the President of Russian Federation for the support of scientific schools NSh-2928.2012.1, and by the Program of the Department of Mathematics of the Russian Academy of Science "Modern problems of theoretical mathematics".

One of the authors (S.K.) gratefully acknowledges being partially supported by the DFG project AL 214/40-1.

References

- [1] A.Yu.Grosberg, A.R.Khokhlov, Giant Molecules: Here, There, and Everywhere, 2nd ed., World Scientific Publishing Company, 2010.
- [2] A.Yu.Grosberg, Disordered polymers, Phys. Usp. (1997) V.40. P.125–158.
- [3] E.I.Shakhnovich, A.M.Gutin, Enumeration of all compact conformations of copolymers with random sequence of links. J. Chem. Phys. 1990. V.93 P.5967-5971.
- [4] Mai Suan Li, D.K. Klimov, D.Thirumalai, Folding in lattice models with side chains, Comput. Phys. Commun. 2002. V.147(1-2), P.625–628, arXiv:cond-mat/0211348v1
- [5] V.S.Pande, A.Yu.Grosberg, T.Tanaka, Thermodynamic procedure to synthesize heteropolymers that can renature to recognize a given target molecule, Proc. Natl. Acad. Sci. USA, 1994, Vol. 91, P. 12976–12979.
- [6] E.I.Shakhnovich, A.M.Gutin, Engineering of stable and fast-folding sequences of model proteins, Proc. Natl. Acad. Sci. USA, 1993, Vol. 90, P. 7195–7199.

- [7] A.V. Chertovich, E.N. Govorun, V.A. Ivanov, P.G. Khalatur, A.R. Khokhlov, Conformation-Dependent Sequence Design: Evolutionary Approach. *Eur. Phys. J. E*, 2004, vol. 13, P. 15–25.
- [8] Yu.F.Krupyanskii, V.I.Goldanskii, Dynamical properties and energy landscape of simple globular proteins. *Phys. Usp.* (2002) V.45. P.1131–1151.
- [9] A.V.Finkelstein, O.B.Ptitsyn, *Protein Physics*. (2002). London–Amsterdam: Academic Press.
- [10] A. Yu. Grosberg, S.K. Nechaev, E. I. Shakhnovich, The role of topological constraints in the kinetics of collapse of macromolecules, *J. Phys. (Paris)*, (1988), V.49 no 12, P.2095–2100.
- [11] V.A.Avetisov, A.Kh.Bikulov, O.A.Vasilyev, S.K.Nechaev, A.V.Chertovich, Some physical applications of random hierarchical matrices, *Journal of Experimental and Theoretical Physics*, 2009, V. 109, No 3, P. 485–504.
- [12] J. Bryngelson, J. Onuchic, N. Socci, P. G. Wolynes, Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis, *Proteins: Structure, Function, and Genetics*, 1995, V.21. P.167–195.
- [13] S. Istrail, F.Lam, Combinatorial algorithms for protein folding in lattice models: a survey of mathematical results, *Communication in information and systems*, 2009, V.9, No. 4, P. 303–346.
- [14] J. Garnier, J.-F. Gibrat, B. Robson, GOR method for predicting protein secondary structure from amino acid sequence, *Methods in Enzymology*, 1996, V. 266, P. 540–553.
- [15] A.N.Nekrasov, Analysis of the information structure of protein sequences: a new method for analyzing the domain organization of proteins. *J. Biomol. Struct. Dyn.* V.21(5), P.615–624, (2004).
- [16] A.N. Nekrasov, A.A. Zinchenko, Structural Features of the Interfaces in Enzyme-Inhibitor Complexes, *Journal of Biomolecular Structure & Dynamics*, 2010, V. 28, No 1. P.85–96.

- [17] Shuangwei Hu, A. Krokhin, A.J. Niemi, Xubiao Peng, Towards Quantitative Classification of Folded Proteins in Terms of Elementary Functions, arXiv:1011.3181